

Expression of the UDP-glucuronosyltransferase 1A Locus in Human Colon

IDENTIFICATION AND CHARACTERIZATION OF THE NOVEL EXTRAHEPATIC UGT1A8*

(Received for publication, October 20, 1997, and in revised form, January 29, 1998)

Christian P. Strassburg‡, Michael P. Manns§, and Robert H. Tukey¶

From the Department of Pharmacology, UCSD Cancer Center, University of California, San Diego, La Jolla, California 92093 and §Department of Gastroenterology and Hepatology, Medizinische Hochschule Hannover, 30625 Hannover, Germany

UDP-glucuronosyltransferases (UGT) catalyze the conjugation of lipophilic exobiotic and endobiotic compounds, which leads to the excretion of hydrophilic glucuronides via bile or urine. By a mechanism of exon sharing, the transcripts of individual first exon cassettes located at the 5' end of the human *UGT1A* locus are spliced to exons 2–5, leading to the expression of at least nine individual UGT genes. Recently, the tissue-specific expression of the *UGT1A* locus has been demonstrated in extrahepatic tissue, leading to the identification of UGT1A7 and UGT1A10 mRNA (Strassburg, C. P., Oldhafer, K., Manns, M. P., and Tukey, R. H. (1997) *Mol. Pharmacol.* 52, 212). However, *UGT1A* expression has not been defined in human colon, which is a metabolically active, external surface organ and a common route of drug administration. UGT1A expression was analyzed in 5 colonic, 16 hepatic, 4 biliary, and 13 gastric human tissue specimens by quantitative duplex reverse transcription-polymerase chain reaction and Western blot analysis, demonstrating lower UGT1A mRNA in the extrahepatic tissues. The precise analysis of unique UGT1A transcripts by exon 1-specific duplex reverse transcription-polymerase chain reaction revealed the expression of UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 in the colon, which are also present in human liver. In addition, the expression of extrahepatic UGT1A10 and UGT1A8 was demonstrated. UGT1A8 was found to be closely related to gastric UGT1A7 with a 93.8% identity of first exon sequences. Expressed UGT1A7 and UGT1A10 protein showed unique catalytic activity profiles, while UGT1A8 was not active with the substrates tested. The ability of UGT1A10 to glucuronidate estrone represents only the second example of a human estrone UGT. The highly related human UGT1A7–1A10 cluster is expressed in a tissue-specific fashion and underlines the role and diversity of physiological glucuronidation at the distal end of the digestive tract.

Glucuronidation represents an important process of metab-

olism capable of transforming lipophilic substrates into hydrophilic glucuronides, a process which facilitates their subsequent elimination via urine, bile, and feces (1). In humans, UDP-glucuronosyltransferases (UGTs)¹ are a key element of conjugation reactions and catalyze the detoxification of a broad array of substrates ranging from bile acids, dietary amines and flavones, phenols, steroids, and the endogenous metabolic product bilirubin (2–7). The *UGT1A* locus² is located on chromosome 2 and potentially encodes nine functional isoforms and three pseudogenes. UGT1A proteins are generated by a strategy of exon sharing (8, 9). Based upon characterization of the *UGT1A* locus, the 3'-donor splice site of the first exons at the 5' end of the locus are spliced to the acceptor splice site of exon 2, thereby combining unique first exons with the constant exons 2–5 at the 3' end of the locus. The presence of TATA-like elements approximately 30 bp upstream of the exon 1 sequences indicates that individual transcriptional regulation may occur with each *UGT1A* gene (8).

Human UGT1A proteins are implicated in the pathophysiology of several disorders. Genetic abnormalities leading to down-regulation or loss of UGT1A function can result in relatively harmless conditions such as Gilbert's or Crigler-Najjar type II disease, but can also constitute a fatal metabolic error present in Crigler-Najjar type I disease (10–13). Additionally, UGT1A proteins have been identified as hepatocellular autoantigens in autoimmune hepatitis type 2 and chronic viral hepatitis D (14, 15), and have been implicated in human carcinogenesis (16). In light of their role in human physiology, hepatic UGT1A proteins have been studied in detail leading to the identification and cloning of UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 from liver tissue (6, 7, 17, 18). However, to understand the complex contribution of the *UGT1A* locus to human xenobiotic and endobiotic metabolism, their regulation in tissues outside the liver, such as the gastrointestinal tract, is required. Recently, the investigation of UGT1A expression in biliary and gastric epithelium has confirmed an individual expression of UGT1A proteins in a tissue specific fashion. This has led to the identification and cloning of the extrahepatic UGT1A7 and UGT1A10 (16, 19). UGT1A7 and UGT1A10 belong to the UGT1A7–10 cluster of *UGT1A* genes that share

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. This work was conducted in part from support from United States Public Health Service Grant GM49135 (to R. H. T.) and the Deutsche Forschungsgemeinschaft Grant Str493/2-1 (to C. P. S.). AF030310.

‡ Recipient of the Pete Lopiccola Fellowship Award, UCSD, for Cancer Research.

¶ To whom correspondence should be addressed: Dept. of Pharmacology, 0636, UCSD Cancer Center, 9500 Gilman Dr., La Jolla, CA 92093-0636. Tel.: 619-822-0288; Fax: 619-822-0363; E-mail: rtukey@ucsd.edu.

¹ The abbreviations used are: UGT, uridine diphosphate-5'-glucuronosyltransferase; *UGT1A*, family 1A uridine diphosphate-5'-glucuronosyltransferase gene locus; RT, reverse transcription; DRT, duplex reverse transcription; PCR, polymerase chain reaction; Sf9, *Spodoptera frugiperda*; LKM-3, liver/kidney microsomal autoantibody type 3; bp, base pair(s).

² A recommended nomenclature for the human *UGT1A* locus, which was outlined during the VIII International Workshop on Glucuronidation and the UDP-glucuronosyltransferases, Iowa City, May 1996, is used throughout this manuscript.

sequence similarities of approximately 90%. Although UGT1A8 expression has not been identified to date, the expression of UGT1A7 and UGT1A10 in extrahepatic tissues and their absence from hepatic tissue suggests that this subset of *UGT1A* genes may have evolved to accommodate nonhepatic glucuronidation requirements of the body. To provide evidence for the role of this cluster of highly related UGT1As in the extrahepatic digestive tract, it was necessary to specifically characterize their tissue distribution.

In this study, *UGT1A* gene expression levels and isoform profiles of liver and the three extrahepatic gastrointestinal tissues, biliary, gastric, and colonic epithelium, were compared. Expression of the *UGT1A* locus in human colon has not been examined previously. The present study analyzed UGT1A mRNA and protein to address the question of whether colon tissue expresses a specific set of UGT1A proteins. The human colon is in prolonged contact with digested matter and is a common route of medical drug application. We have previously examined the expression of the *UGT1A* locus in liver and biliary epithelium, and expand on this analysis by comparing expression in colonic epithelium. Evidence is provided for a tissue specific gene expression of the *UGT1A* locus in colon, with the characterization of a novel extrahepatic UGT1A8 protein.

MATERIALS AND METHODS

Hepatic and Extrahepatic Tissue Samples

Liver tissue samples were obtained from 16 patients undergoing hemihepatectomy or liver transplantation for hepatocellular carcinoma, and biliary epithelium was obtained from four healthy gallbladder specimens as reported previously (19). Gastric tissue was obtained from 13 patients undergoing three fourths gastrectomy for gastric carcinoma, and colonic tissue was collected from five surgical specimens removed during hemicolectomy for adenocarcinoma. All surgical specimens were obtained from the University of Hannover Medical Center, Hannover, Germany. Only healthy tissue that did not display signs of necrosis or any pathoanatomical abnormalities was included. All samples were immediately snap frozen in liquid nitrogen and continuously stored at -80°C until analysis.

RNA Isolation

Approximately 200 mg of frozen tissue were pulverized in a mortar filled with liquid nitrogen. Tissue powder was immediately lysed in acidic phenol/guanidinium isothiocyanate solution (Tripure, Boehringer Mannheim, Germany), and RNA was extracted as reported previously (19). Concentrations were measured by optical density at 260 and 280 nm, and RNA samples were stored in water at -80°C until further use.

cDNA Synthesis and Documentation of DNA Free RNA

Three μg of total RNA were denatured for 10 min at 70°C in the presence of 0.5 μg of oligo(dT) primer and placed on ice. The reaction conditions were adjusted to 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2.5 mM MgCl_2 , 10 mM dithiothreitol, 0.5 mM of each dNTP in a total volume of 19 μl , which was incubated at 42°C for 5 min. Synthesis was started by adding 200 units of reverse transcriptase (Superscript II, Life Sciences, Inc.). The final volume of 20 μl was incubated at 42°C for 50 min, at 70°C for 15 min, and placed on ice.

RT-PCR was used to exclude contamination of RNA with DNA. Amplification of human β -actin with the primers 5'-ggcggcaccaccatgtaccct-3' and 5'-aggggccggactcgtcactac-3' (20) spanning the exon 4/intron 5/exon 5 junction of human β -actin leads to a 202-bp product when a cDNA template is present, but results in a 312-bp product when a genomic DNA template is present. Both products can be clearly distinguished by 2% agarose electrophoresis. All cDNAs used displayed a 202-bp product and were thus free of contaminating genomic DNA.

Duplex Reverse Transcription Polymerase Chain Reaction (DRT-PCR)

All UGT1A Transcripts—The presence of UGT1A transcripts in total tissue RNA was detected by the amplification of a 487-bp fragment from the conserved UGT1A exons 2–5 using DRT-PCR as previously reported in detail (19). In brief, amplification with UGT1A primers at concen-

trations of 2 μM each was first carried out for six cycles using a Perkin-Elmer 480 Thermocycler (all PCR reactions). After the addition of β -actin primers to a concentration of 0.4 μM , cycling was continued for a total of 32 cycles at 94°C for 1 min, 59°C for 1 min, and 72°C for 1 min. PCR was preceded by a hot start of 3 min at 94°C and followed by a 7-min elongation at 72°C . DRT-PCR products were separated in a 2% agarose gel stained with ethidium bromide. Polaroid (Cambridge, MA) type 665 positive/negative film was used to quantify bands by laser densitometry on a LKB 2222–020 UltraScan XL densitometer (LKB, Bromma, Sweden). Arbitrary units were calculated relative to β -actin products. Independent and combined linear kinetics for both products were established as described previously (16, 19). Statistical analysis was performed using Student's *t* test.

Exon 1-specific DRT-PCR—The *UGT1A* locus predicts the existence of nine proteins termed UGT1A1 and UGT1A3–1A10. UGT1A2, UGT1A11 and UGT1A12 lack an uninterrupted open reading frame and have therefore been identified to be pseudogenes (8). DRT-PCR detection of all nine UGT1A transcripts predicted by the human *UGT1A* locus was performed using nine exon 1-specific sense primers and two antisense primers located within exons 2–5 or within a common portion of the 3' end of the first exons (19). Primer sequences can also be obtained at <http://www.molpharm.org/cqi/content/full/52/2/212>. As already reported elsewhere (19), exon-specific primers were generated which lead to RT-PCR products of distinct molecular sizes: UGT1A1, 644 bp; UGT1A3, 483 bp; UGT1A4, 572 bp; UGT1A5, 659 bp; UGT1A6, 562 bp; UGT1A7, 754 bp; UGT1A8, 514 bp; UGT1A9, 392 bp; and UGT1A10, 478 bp. Coamplification of the UGT1A first exon and β -actin sequences was performed using three cycling protocols, UGT1A1 and UGT1A6, 94°C (1 min), 59°C (1 min), 72°C (1 min); UGT1A3, UGT1A4, UGT1A5, 94°C (1 min), 56°C (1 min), 72°C (1 min); and UGT1A7, UGT1A8, UGT1A9, UGT1A10, 94°C (1 min), 64°C (1 min), 72°C (1 min). Each protocol was preceded by a 3-min incubation of the reaction mixture at 94°C and followed by a 7-min elongation at 72°C . The specificity and kinetics of this assay have previously been documented in detail (16, 19). Experiments were performed in duplicate, and controls without cDNA, primers, or thermophilic polymerase were included. Quantification of products by laser densitometry was performed as described above.

Isolation of Microsomal Protein from Hepatic and Extrahepatic Gastrointestinal Tissue

Liver, gastric, gall bladder, and colonic tissue samples of approximately 150 mg were pulverized under liquid nitrogen, resuspended in 1 ml of buffer (50 mM Tris-HCl, pH 7.4, 10 mM MgCl_2) and homogenized with a Potter-Elvehjem tissue grinder. The tissue homogenate was centrifuged at $10,000 \times g$ for 5 min at 4°C , and the supernatant was collected. The pellet was resuspended in 0.5 ml of buffer and centrifuged, and the supernatant was collected. The combined supernatants were centrifuged at $150,000 \times g$ for 60 min at 4°C , and the pellet was resuspended in 0.2 mL of buffer. Protein concentration was determined by the method of Bradford (21). Microsomal protein was stored at -80°C .

Western Blot Analysis

Microsomal protein from human liver, gastric, biliary, and colonic epithelium was boiled for 90 s in loading buffer (2% sodium dodecyl sulfate, 62.5 mmol/liter Tris-HCl, pH 6.8, 10% glycerol, and 0.001% bromophenol blue) with β -mercaptoethanol and resolved by 8% sodium dodecyl sulfate-polyacrylamide gel electrophoresis prior to electrotransfer onto a nitrocellulose membrane. Also included was a 20- μg sample of total *Spodoptera frugiperda* (Sf9) cell lysate expressing recombinant UGT1A1 protein (15). Immunodetection was performed following published protocols (15). UGT1A protein was detected using a human anti-UGT1A antiserum of a patient with chronic hepatitis D (liver/kidney microsomal autoantibody type 3 (LKM-3)-positive serum) diluted 1:1000. This autoimmune serum has been characterized to specifically recognize UGT1A protein (14). Visualization was achieved with an alkaline phosphatase-conjugated goat anti-human IgG diluted 1:1000.

Rabbit Anti-UGT1A Antiserum

Eight copies of the peptide SSLHKDRPVEPLDLA were synthesized on a branched lysine multiple antigen peptide core (Research Genetics Inc., Huntsville, AL). This 15-mer sequence is located between amino acids 441 and 455 on exon 5 of the constant carboxyl-terminal portion of the UGT1A7 sequence and is common to all UGT1A proteins. Animals were injected subcutaneously into dorsal sites with 0.5 mg of peptide on

day 1, 14, 42, and 56, and serum was collected after day 70. Antibody production was monitored by enzyme-linked immunosorbent assay with 1 μ g of peptide used as a solid phase per well. Titers at 70 days exceeded 1:204,800. Western blot analysis of recombinant baculovirus generated UGT1A1, UGT1A7, UGT1A8, and UGT1A10 protein was performed using a 1:1500 dilution of rabbit anti-UGT1A immunoglobulin as first antibody and a goat anti-rabbit alkaline phosphatase conjugated IgG diluted 1:4500 as the visualization antibody.

Cloning and Characterization of Human Colonic UGT1A8

cDNA was synthesized from 3 μ g of total colonic RNA as described above. For amplification of the entire UGT1A8 open reading frame, two primers were generated by automated phosphoramidite chemistry (UCSD Cancer Center Oligonucleotide Core Facility). Primer 1, 5'-*tcgggatccagttctctcatg(start)gctgcac-3'*, was designed to incorporate a *Bam*HI restriction site upstream of the ATG start codon, and primer 2, 5'-*cacaagcttctca(stop)atgggtcttggattgtgg-3'*, was designed to introduce a *Hind*III restriction site downstream of the inherent TGA-stop codon to facilitate directional cloning of the product. In a reaction volume of 100 μ l of 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 2.75 mM MgCl₂, 0.1% Triton X-100, 0.3 mM each dNTP, 1 μ M of each primer, and containing 5 units of *Taq* DNA polymerase (Promega, Madison, WI) cycling was carried out at 94 °C (60 s), 58 °C (60 s), and 72 °C (80 s) for 35 cycles. The protocol was preceded by an incubation of 5 min at 94 °C and followed by an extended elongation time of 7 min at 72 °C. The amplified product of approximately 1600 bp was digested with *Bam*HI and *Hind*III (New England Biolabs), gel-purified (Quiagen, Hilden, Germany), and inserted into pBluescript KS+ (Stratagene, La Jolla, CA). The sequence of 1593 bp was determined by dideoxy sequencing (Life Sciences, Cleveland, OH) on both strands using internal oligonucleotide primers.

Expression of Recombinant UGT1A7, UGT1A8, and UGT1A10 Protein in Baculovirus

The entire open reading frame of the previously described UGT1A7 and UGT1A10 cDNAs (19) and the colon-derived UGT1A8 cDNA were inserted into a pBluebac 4 (Invitrogen, San Diego, CA) baculovirus transfer vector and transfected into log phase insect cells (Sf9) which were cultured in monolayers as described previously (15, 22). UGT1A7, UGT1A8, and UGT1A10 recombinant virus stock with a titer of $1-2 \times 10^8$ plaque-forming units/ml was generated and subsequently used to infect log phase Sf9 cell monolayers at a multiplicity of infection of 10 together with previously reported UGT1A1 recombinant virus as a control (15). Cells were collected at 60 h postinfection and lysed with a Potter-Elvehjem tissue grinder in buffer containing 50 mM Tris-HCl, pH 7.4, 10 mM MgCl₂. Recombinant UGT cell lysates were stored at -70 °C until analysis. UGT1A7, UGT1A8, and UGT1A10 were additionally cloned into pcDNA3.1 (Invitrogen) and pSVL (Pharmacia Biotech Inc.) vectors and expressed in COS-1 cell cultures following published protocols (23).

Catalytic Activity Assay of Human Extrahepatic UGT1A7, UGT1A8, and UGT1A10

Glucuronidation substrates (all 46 tested substrates are listed in Tables I and II) were solubilized in methanol with the exception of 7-hydroxybenzo(a)pyrene, which was resuspended in acetone (7-hydroxybenzo(a)pyrene, and 8-hydroxybenzo(a)pyrene were obtained from the National Cancer Institute Chemical Carcinogen Repository, Midwest Research Institute, Kansas City, MO, all other substrates were from Sigma-Aldrich, St. Louis, MO). Catalytic activities of 100 μ g of Sf9 cell or COS-1 cell lysate containing recombinant UGT1A7, UGT1A8, or UGT1A10 were assayed in duplicate in the presence of 0.1 mM UDP-glucuronic acid, 0.08 mg/ml phosphatidylcholine, 0.04 μ Ci of ¹⁴C-labeled glucuronic acid, 100 μ M test substrate in 50 mM Tris-HCl pH 7.6, and 10 mM MgCl₂ for 60 min at 37 °C. Protein was subsequently precipitated by the addition of ethanol followed by centrifugation. Lyophilized supernatants were resuspended in 50 μ l of methanol and separated by thin layer liquid chromatography in *n*-butanol/acetone/acidic acid/water (35:35:20:20%). ¹⁴C-labeled glucuronides were detected by autoradiography. Glucuronide bands were then removed from the TLC plates, counts were determined, and specific catalytic activities were calculated in picomoles of glucuronide formed/mg of protein/min.

RESULTS

UGT1A Expression in Human Hepatic and Extrahepatic Gastrointestinal Epithelium—All human *UGT1A* gene transcripts display a unique 5' terminus characteristic of each of the individual transferase and a common 3' portion encoded by exons 2–5. The 3' region is identical in all members encoded by

the *UGT1A* locus and can therefore be exploited to analyze overall *UGT1A* expression (Fig. 1). DRT-PCR amplifying a 487-bp fragment of the *UGT1A* constant region was used to analyze the presence of *UGT1A* mRNAs and their relative levels in a panel of 16 normal human liver RNAs in comparison to gastric ($n = 13$), biliary ($n = 4$), and colonic ($n = 5$) epithelial tissue. *UGT1A* expression was detected in all samples studied. Quantification of DRT-PCR products by laser densitometry, however, revealed significant differences of steady state levels between hepatic and extrahepatic tissues. The highest level of *UGT1A* expression was detected in hepatocellular tissue, which was significantly higher than in biliary ($p = 0.001$), gastric ($p = 0.00001$), or colonic ($p = 0.0009$) epithelium. The comparisons of *UGT1A* expression between biliary and gastric ($p = 0.3$), biliary and colonic ($p = 0.1$), and gastric and colonic tissue ($p = 0.6$) were not significant.

Hepatic and Extrahepatic UGT1A Protein Expression Is Heterogeneous—*UGT1A* expression in endoplasmic reticulum prepared from human hepatocellular, biliary, gastric, and colonic tissue was analyzed by Western blot. Monospecific immunoglobulin capable of detecting the individual *UGT1A* proteins is not available. However, to analyze total hepatic and extrahepatic *UGT1A* protein expression, a human LKM-3 anti-*UGT1A* autoimmune antibody was employed with epitope identity directed specifically against *UGT1A* proteins (14, 15). Human liver *UGT1A* expression as detected by anti-*UGT1A* immunoglobulin Western blot analysis displayed a distinctly different pattern when compared with the other three extrahepatic microsomal preparations (Fig. 2). In hepatic microsomes, several *UGT1A* subspecies were detected around the typical molecular weight of 55 kDa. In gastric, biliary, and colonic epithelial microsomes, the pattern was more uniform and clearly different from that found in liver. This indicates the presence of different *UGT1A* proteins expressed in liver and confirms a qualitative and tissue-specific differential expression of the human *UGT1A* locus. However, defining the specificity of the *UGT1A* expression with antibodies is difficult since little difference in the resolution of the different proteins was apparent from Western blot analysis. In the absence of specific *UGT1A* immunoglobulin to detect the individual proteins, experiments were designed to characterize specific *UGT1A* gene expression at the mRNA level.

Tissue-specific Expression of the UGT1A Locus in the Human Gastrointestinal Tract Identifies a Novel Transcript Localized to Colonic Epithelium—Analysis of *UGT1A* gene expression was performed at the mRNA level employing *UGT1A* exon 1-specific DRT-PCR. As reported previously, human hepatic, gastric, and biliary tissues are characterized by a unique and tissue-specific differential expression of the *UGT1A* locus, which is demonstrated in Fig. 3, as previously reported (19). In contrast to hepatocellular epithelium, both gastric and biliary epithelium express *UGT1A10*, while gastric epithelium also expresses *UGT1A7*. Within these three tissues *UGT1A9* expression was found exclusively in the liver. In human colon, however, a unique expression pattern was identified, which is characterized by the combined expression of *UGT1A9* and *UGT1A10* in addition to the unique expression of *UGT1A8*. In contrast to gastric tissue, *UGT1A7* transcripts were not detected in colon. Interestingly, the expression of *UGT1A10* is localized to extrahepatic tissues but is found to be absent in liver. These data demonstrate precise tissue-specific expression of genes encoding the *UGT1A7-1A10* cluster of transferases, and show that the *UGT1A* locus is expressed in the human extrahepatic gastrointestinal tract. In addition, the identification of *UGT1A8* represents the first example of expression of this gene in any human tissue.

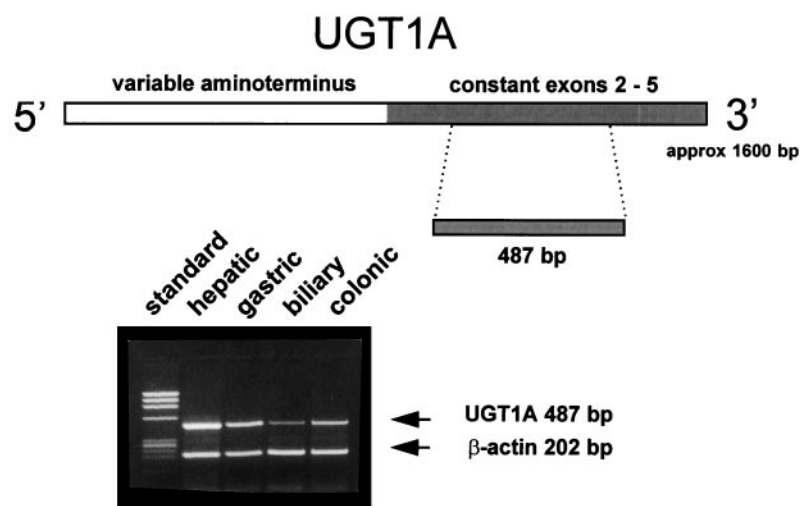


FIG. 1. Total UGT1A expression analyzed by quantitative DRT-PCR. A 487-bp fragment of the constant carboxyl terminus which encodes exons 2–4 was amplified by DRT-PCR. The *top panel* is a schematic of a UGT1A transcript and indicates the location of the amplified fragment. Co-amplification of the 487-bp UGT1A and 202-bp human β -actin PCR products separated in a 2% agarose gel stained with ethidium bromide is demonstrated using hepatic, gastric, biliary, and colonic cDNA. One μ g of *Hae*III-digested ϕ X174RF DNA was included in the first lane as a standard. The *bottom panel* summarizes the relative levels of total UGT1A expression in 16 hepatic, 13 gastric, 4 biliary, and 5 colonic tissue samples. The *asterisk* (*) indicates statistically significant comparisons ($p < 0.05$) to hepatic expression using the Student's *t* test.

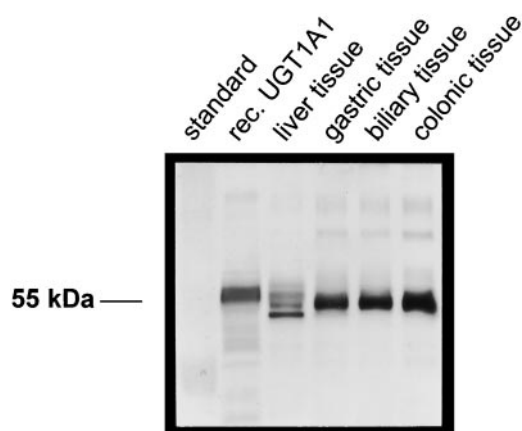
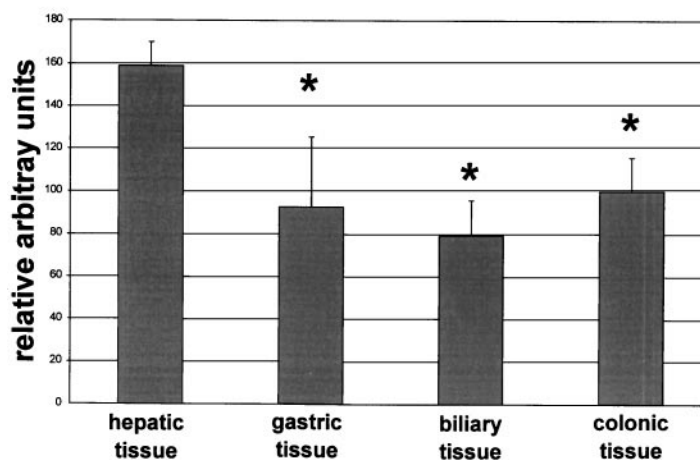


FIG. 2. Western blot analysis of differences between hepatic and extrahepatic microsomal UGT1A protein. Microsomal protein prepared from human liver, gastric, biliary, and colonic tissue were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis and blotted onto a nitrocellulose membrane. Also included was a sample of Sf9 cell microsomes prepared from cells that express recombinant UGT1A1 (*rec. UGT1A1*). Immunodetection of the UGTs was accomplished with a human anti-UGT1A LKM-3 autoimmune serum as described under “Materials and Methods.”

Characterization of Colonic UGT1A8—UGT1A8 was cloned from colon tissue RNA and characterized by DNA sequence analysis. The UGT1A8 cDNA displayed an open reading frame of 1593 bp and encoded a protein of 530 amino acids, with a carboxyl-terminal portion of 246 amino acids that was identical

to exons 2–5 of the other UGT1A cDNAs (Fig. 4). Divergence in amino acid sequence with other UGT1A proteins are located within the amino-terminal 284 residues, corresponding to the variable first exon. The UGT1A7–1A10 proteins display a high degree of similarity. UGT1A8 shares 93.8% identity with UGT1A7, 93.7% with UGT1A9, and 90.2% with UGT1A10. The reduced level of similarity between UGT1A8 and UGT1A10 is largely due to the classical signal peptide sequence located at the 5' end of all known UGTs which is lacking in UGT1A10 (19, 24). Alignment of UGT1A7–1A10 (Table I) demonstrates that the extrahepatic UGT1A7 and UGT1A8 are closely related and that divergence is greater with UGT1A9 and UGT1A10 (Fig. 5).

Catalytic Activities of Extrahepatic UGT1A7, UGT1A8, and UGT1A10—Western blot analysis of expressed protein in Sf9 cells demonstrated the generation of recombinant UGT1A7, UGT1A8, and UGT1A10 protein were at comparable levels (Fig. 6). Previously reported recombinant human UGT1A1 and liver microsomal protein served as positive controls for the identification of UGT1A protein. All three UGT1A proteins displayed striking differences of catalytic activities. UGT1A7 was capable of forming glucuronides with a large number of planar (*e.g.* 4-nitrophenol) and complex (*e.g.* 4-*tert*-butylphenol, propylgallate) phenols in addition to substrates with coumarin, flavone, anthraquinone, steroid hormone, and naphthol structures, and with hydroxylated benzo(a)pyrenes (Fig. 7 and Tables II and III). In contrast, UGT1A10 protein, while also glucuronidating planar and phenolic substrates, displayed specificity for steroid hormone substrates such as estrone, androsterone, β -estradiol, and 4-hydroxyestrone. These data rep-

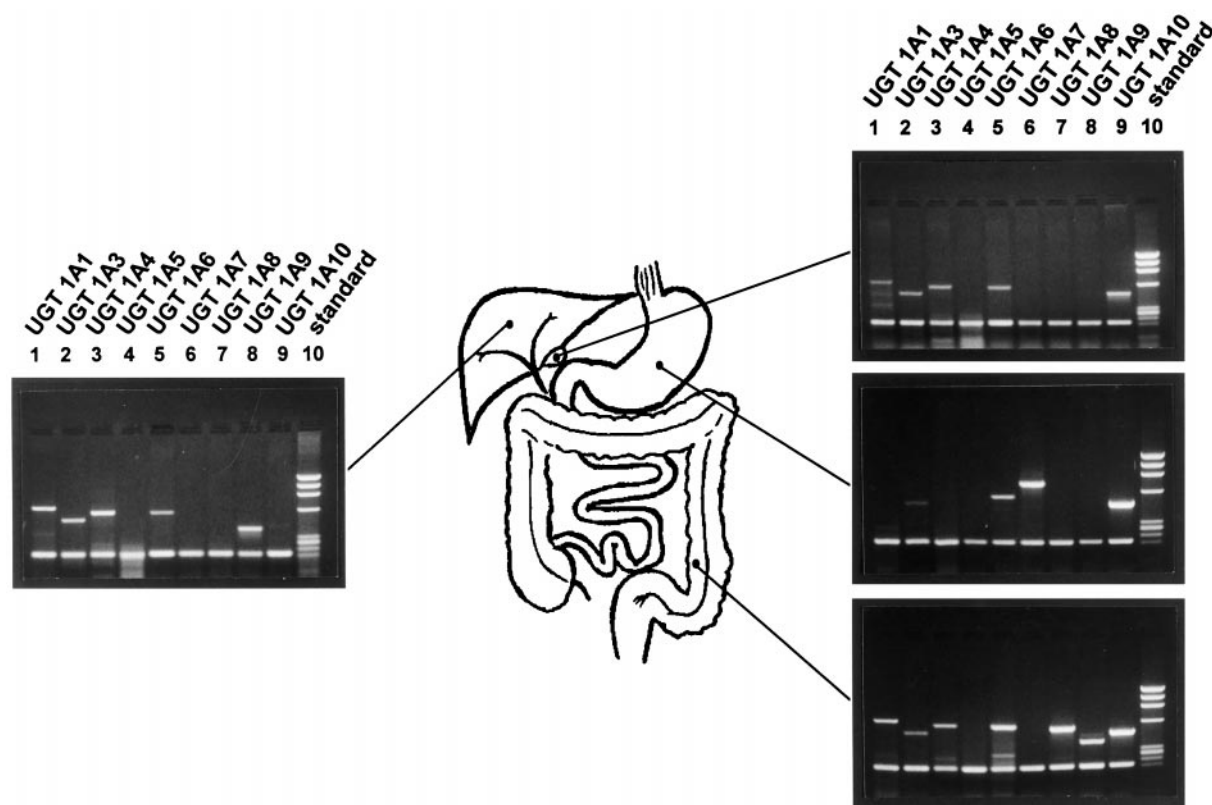


FIG. 3. **Differential regulation of the UGT1A locus in the human gastrointestinal tract (schematic).** Products of exon 1-specific DRT-PCR specifically detecting transcripts of all known UGT1A isoforms were separated in 2% agarose and stained with ethidium bromide. The left panel and the top two panels on the right demonstrate the differential regulation of the *UGT1A* locus in hepatic, gastric, and biliary epithelium. The lower panel on the left shows the specific UGT1A regulation found in colonic epithelium. Colon tissue is characterized by the expression of UGT1A10, which is found in all other extrahepatic tissues (three panels on the right), but not in liver (left panel). In addition, colonic tissue is characterized by the unique expression of UGT1A8, a transcript that is not found in other extrahepatic tissues. UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 are transcribed both in hepatic and colonic tissue. Note that the expression of UGT1A7 is unique to gastric tissue (middle panel on the right). Standard, *Hae*III-digested ϕ X174RF DNA.

resent the first report of specificity for phenolic substrates (UGT1A7 and UGT1A10) and steroid hormones (UGT1A10) for these extrahepatic UGT1A proteins. The capability of UGT1A7 to glucuronidate complex phenols such as 4-*tert*-butylphenol in addition to planar phenols and flavones is similar to the profiles already reported for UGT1A9 (25). Despite its high degree of sequence similarity to UGT1A7, UGT1A8 displays none of the activities documented here for UGT1A7 or for UGT1A10. These results were confirmed by the additional expression of UGT1A7, UGT1A8, and UGT1A10 in COS-1 cells and subsequent catalytic activity assays. In both Sf9 cell and COS-1 cell systems UGT1A7 and UGT1A10 conjugated the same panel of substrates, but UGT1A8 protein was not capable of glucuronidating any of the substrates in this series (not shown).

DISCUSSION

Glucuronidation plays a central role in the metabolism of endobiotic and exobiotic compounds. In humans, UGTs have evolved into a superfamily of proteins that have been identified as the UGT1 and UGT2 families based upon sequence similarities (26). The *UGT1A* gene locus is linked to a number of human disorders that emphasize the physiological importance of these proteins. In Crigler-Najjar type I disease the low or absent activity of functional UGT1A1 impairs the human body's only effective means of detoxifying and eliminating bilirubin (11–13). This fatal metabolic error consequently leads to hyperbilirubinemia and kernicterus (10). Other than in this example, most proteins encoded by the human *UGT1A* locus are characterized by a broad overlap of substrate specificities (3, 4, 8, 25). The high degree of sequence similarity and the

unavailability of diagnostic substrates for all UGT1A proteins has limited the characterization and identification of additional members of the UGT1A family.

This study compares *UGT1A* regulation in the liver and three extrahepatic tissues. Gastric and colonic epithelium represent the two physiologically most important extrahepatic tissues of the gastrointestinal tract. They are both external surface tissues with extensive contact to ingested exobiotic compounds, are metabolically active, and are characterized as leading locations for chemically induced tumorigenesis. In drug therapy, gastric and colonic mucosa are frequently the first site of contact, and their cellular metabolic capacities may have a profound impact on drug efficacy and absorption. The anatomy of the gastrointestinal tract suggests that gastrointestinal tissues would be required to exert a refined control of detoxifying and drug metabolizing enzymes, and that this control may be specific from organ to organ, and from the proximal to the distal portions of the digestive tract.

Western blot analysis demonstrated the expression of UGT1A protein in liver, stomach, biliary epithelium, and colon. Analysis of total UGT1A expression by quantitative DRT-PCR provided evidence for a higher level of UGT1A mRNA in liver than in three extrahepatic digestive tract tissues. In addition, UGT1A mRNA levels between gastric, biliary, and colonic epithelial samples did not differ significantly. These data indicate a quantitative differential expression of the human *UGT1A* locus between hepatic and the other extrahepatic tissues. Interestingly, human colon appeared to have levels of UGT1A comparable to stomach and biliary epithelium. This implicates

FIG. 4. Sequence alignment comparing the three extrahepatic UGT1A proteins UGT1A7, UGT1A8, and UGT1A10. All three proteins share a common carboxyl terminus of 245 residues, which is encoded by the common exons 2–5. The location of the proteins that are encoded by exon 1 are marked by the arrow. All differences between the proteins reside within the 285 amino-terminal residues. The complete sequence of the colonic UGT1A8 is listed while only differences in UGT1A7 and UGT1A10 are shown. The shaded box indicates the predicted signal sequence portion calculated by the method of von Heijne (24). Both UGT1A7 and UGT1A8 display a leader sequence typical of a classical signal peptide and is shown in the shaded area. The similar region of UGT1A10 does not represent a consensus signal peptide. (The novel UGT1A8 sequence has been submitted to GenBank™ as accession no. AF030310.)

UGT1A8	MARTGWTSPIPL-CVSL--LLTCGFAEAGKLLVVPMDGSHWFTMQSVVE	46
UGT1A7	RA GLL -Y C -- K	46
UGT1A10	PRRDQ RSFM TAD WL --- A	46
UGT1A8	KLILRGHEVVVVVPEVSWQLGKSLNCTVKTYSTSYTLEDLDFREFMDFADA	96
UGT1A7	R Q V	96
UGT1A10	ER QN V H	96
UGT1A8	QWKAQVRSLSFLSLSNGFFNLFSSHCRLFNDRRLVEYLKESFFDAVF	146
UGT1A7	R T PL A LT I D N K C	146
UGT1A10	AQ I LM S LD K	146
UGT1A8	LDPFDACALIVAKYFSLPSVVFARGIGCHYLEEGAQC PAPLSYVPRILLG	196
UGT1A7	G F L	196
UGT1A10	T G T F RH ND	196
UGT1A8	FSDAMAFKERVNRHILHLEHLFCQYFSKNALEIASEILQTPVTAYDLYS	246
UGT1A7	T W M P F V	246
UGT1A10	T W V D LLR	246
UGT1A8	HTSIWLLRTRDFVLDYPPKVPMPNMFIFIGGINCHQ GKPLPMEFEAYINASGE	296
UGT1A7	E V	296
UGT1A10		296
UGT1A8	HGIVVFLGSMVSEIPEKKAMAADALGKIPQTVLWRYTGTGPSNLANT	346
UGT1A7		
UGT1A10		
UGT1A8	ILVKWLPQNDLLGHPMTRAFITHAGSHGVYESICNGVPMVMPLFGDQMD	396
UGT1A7		
UGT1A10		
UGT1A8	NAKRMETKGAGVTLNVLEMTSEDLLENALKAVINDKSYKENIMRSLSLHKD	446
UGT1A7		
UGT1A10		
UGT1A8	RPVEPLDLAVFWVFVMRHKGAPHLRPAAHDLTWYQYHSLDVI GFLAVV	496
UGT1A7		
UGT1A10		
UGT1A8	LTVAFITFKCCAYGYRCKLGKGRVKKAHKSKTH	530
UGT1A7		
UGT1A10		

TABLE I
Similarity of the UGT1A7, UGT1A8,
UGT1A9, and UGT1A10 proteins

Listed are the number of identical amino acids between the compared UGT1A proteins (19, 17). Similarity decreases from UGT1A7 to UGT1A10 (compare Fig. 5).

	UGT1A7	UGT1A8	UGT1A9	UGT1A10
UGT1A7				
UGT1A8	497			
UGT1A9	487	491		
UGT1A10	477	479	466	

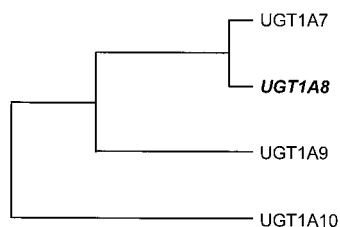


FIG. 5. Dendrogram of the cluster of UGT1A7, UGT1A8, UGT1A9, and UGT1A10. The alignments were conducted with the full protein sequences. All four UGTs are regulated in extrahepatic tissue and share an overall similarity of 85.2%. UGT1A7 and UGT1A8, both of which are uniquely expressed in gastric and colonic epithelium, respectively, are most closely related. Divergence increases toward UGT1A9 and UGT1A10 and may be indicative of the evolutionary development of these proteins. (Compare Table I.)

the colon as an active organ of extrahepatic glucuronidation.

Recently, UGT1A7 and UGT1A10 transcripts have been identified in gastric and biliary tissues (19). This has conclu-

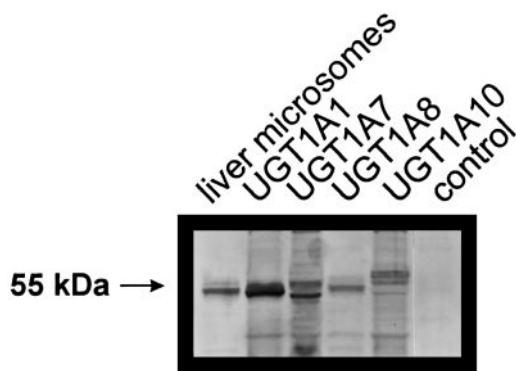


FIG. 6. Western blot of baculovirus expressed extrahepatic UGT1A7, UGT1A8 and UGT1A10. Twenty-five μ g of total Sf9 cellular homogenate expressing UGT1A7, UGT1A8, and UGT1A10 were loaded as indicated. Included in the Western blot are a sample of human liver microsomes (10 μ g) and Sf9 cell microsomes (15 μ g) from cells expressing the human UGT1A1 protein (15). Immunodetection of the expressed UGT proteins was accomplished with a rabbit anti-UGT1A antiserum, as outlined under "Materials and Methods."

sively demonstrated a qualitative difference of UGT1A expression in extrahepatic and hepatic tissues, but also between different extrahepatic tissues. To examine UGT1A expression differences at the protein level, tissue microsomal protein was analyzed by Western blot. Using a previously characterized human anti-UGT1A antiserum (anti-LKM-3, positive) from a patient with chronic hepatitis D virus infection (14), immunodetection identified different patterns of UGT1A protein in hepatic *versus* extrahepatic microsomes. In liver microsomes, a

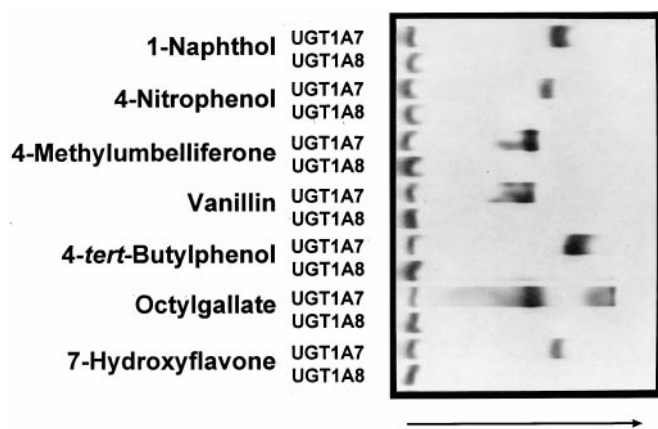


FIG. 7. Autoradiography showing ^{14}C -labeled glucuronides synthesized using UGT1A7 and UGT1A8 protein expressed in Sf9 cells. For these assays, 100 μg of total cellular protein were used, as outlined under "Materials and Methods." Shown is the thin layer liquid chromatography profile of the glucuronides. Recombinant UGT1A7 catalyzed the glucuronidation of all shown substrates, while UGT1A8 was not catalytically active toward any of the tested substrates despite a sequence similarity with UGT1A7 of 93.8%. Arrow, direction of thin layer liquid chromatography.

TABLE II
Differential catalytic activities of gastric-expressed UGT1A7, colon-expressed UGT1A8, and extrahepatic UGT1A10

One hundred μg of baculovirus expressed UGT1A7, UGT1A8, and UGT1A10 were used in each assay with substrate concentrations at 100 μM . ^{14}C -Labeled glucuronides separated by thin layer liquid chromatography were visualized on x-ray film and glucuronide bands were removed for liquid scintillation counting. Specific activity rates were calculated from duplicate experiments in pmol of glucuronide formed/mg of protein/min. Expression of UGT1A7, UGT1A8, and UGT1A10 in COS-1 cells confirmed the expression pattern demonstrated in the table (not shown).

Substrate	UGT1A7	UGT1A8	UGT1A10
100 μM	pmol/mg/min		
Phenolic structures			
4-Nitrophenol	28.1	ND ^a	10.2
2-Hydroxybiphenyl	12.4	ND	3.8
4-Hydroxybiphenyl	14.6	ND	49.3
4-Isopropylphenol	41.0	ND	18.4
4- <i>tert</i> -Butylphenol	123.7	ND	4.4
Propyl gallate	119.4	ND	45.1
Octyl gallate	175.2	ND	33.3
Vanillin	94.9	ND	87.6
Anthraquinone and flavone structures			
1,8-Dihydroxyanthraquinone	17.2	ND	6.3
7-Hydroxyflavone	57.2	ND	34.9
Coumarin structures			
4-Hydroxycoumarin	4.1	ND	4.3
4-Methylumbelliferone	219.5	ND	10.2
Naphthol structures			
1-Naphthol	15.5	ND	59.9
2-Naphthol	33.2	ND	10.3
Carcinogens			
7-Hydroxybenzo(<i>a</i>)pyrene	42.7	ND	14.4
8-Hydroxybenzo(<i>a</i>)pyrene	3.5	ND	3.2
Steroid hormone structures			
Estradiol	6.3	ND	18.2
2-Hydroxyestradiol	1.5	ND	9.5
Androsterone	ND	ND	4.2
β -Estradiol	ND	ND	47.9
4-hydroxyestrone	ND	ND	15.1
Estrone	ND	ND	10.7

^a ND, not detected.

unique banding of proteins was observed, contrasting the more uniform banding pattern observed in gastric, biliary, and colonic microsomes. Interestingly, Western blot analysis identifies proteins and displays a banding pattern quite different from that of UGT1A expression in extrahepatic tissue. Since we

TABLE III
Substrates not glucuronidated by UGT1A7, UGT1A8, or UGT1A10

1,2-Benzanthracene
1,2-Dihydroxy-diphenylamine
17 α -Hydroxyprogesterone
2-Aminophenol
2-Hydroxyestriol
2-Hydroxytyramine
4-Acetamidophenol
4-Aminobiphenyl
6-Hydroxymelatonin
α -Naphtholflavone
α -Naphthylamine
All- <i>trans</i> -retinoic acid
Amitriptyllin
β -Naphtholflavone
<i>cis</i> -Retinoic acid
Desipramine
Diphenhydramine
DL-Tryptophan
Hydroquinone
Hyodeoxycholic acid
Ibuprofen
Imipramine
<i>p</i> -Hydroxybenzoic acid
Testosterone

are capable of identifying expression of the UGT1A locus in liver and colon, it is quite possible that the lower migrating protein in liver (Fig. 2) is a unique UGT that has not been identified.

Colon tissue displayed a unique expression of the UGT1A locus that differed from the other three studied tissues. The colon samples were found to express UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9, all of which were present in liver. In addition, colon was characterized by the presence of UGT1A10, which is expressed in all extrahepatic digestive tract tissues (19). Interestingly, colon is unique in the expression of UGT1A8. The expression of seven individual UGT1A genes in colon is the most diverse expression pattern of the UGT1A locus so far identified in the human gastrointestinal tract. It exceeds the expression of the four isoforms in gastric, five in biliary, and five in hepatic epithelium. This finding suggests that, although UGT1A expression is more abundant in liver, the greater diversity in UGT1A locus expression in human colon may reflect glucuronidation and drug-metabolizing requirements necessary at the distal end of the gastrointestinal tract. The long exposure time of colon mucosa to intraluminal compounds, fermentation products of the indigenous flora, and the collection of substances that have escaped detoxification in the proximal digestive system may have led to the requirement to express these individual UGT1A proteins. On the other hand, a unique tissue specificity is maintained through the expression of UGT1A8. In the panel of tissues analyzed, UGT1A7 and UGT1A8 appear to be transcripts that are highly tissue-specific, while UGT1A10 is a transcript that discriminates hepatic versus nonhepatic epithelium.

The colon-derived UGT1A8 cDNA of 1593 bp was isolated and found to be structurally similar to other UGT1A transcripts. Protein sequence alignments with UGT1A7 and UGT1A10 show that all differences are located within the first 284 amino acids, a region that corresponds to the divergent exon 1 found in other UGT1A transcripts. The carboxyl-terminal 246 amino acids are identical in all three sequences and correspond to the constant exons 2–5 (2). To provide evidence for a physiological relevance of these extrahepatic UGT proteins, UGT1A7, UGT1A8, and UGT1A10 were expressed in baculovirus and their catalytic activities characterized. While the UGT1A7–1A10 genes share approximately 90% sequence similarity, their catalytic specificities were found to be remark-

ably divergent. From the panel of 46 substrates used, UGT1A7 is capable of glucuronidating a wide variety of phenolic compounds and flavones, anthraquinones, coumarins, and carcinogens. In addition, UGT1A10 was found to be an active protein despite its characteristic absence of a classical signal peptide sequence. Contrasting the activity profile of UGT1A7 however, UGT1A10 catalyzes the glucuronidation of a number of steroid hormones in addition to planar and complex phenolics. This finding represents only the second example of human estrone UGT activity (6). UGT1A8, however, did not catalyze the glucuronidation of the series of substrates tested. These findings indicate a very different activity profile for UGT1A8, which may be considerably focused and specialized. This appears to be supported by its coexpression in colon with UGT1A9. UGT1A9 has been previously characterized to glucuronidate a broad variety of phenolic substrates, flavones, and anthraquinones, including those found in this study for UGT1A7 (25) (Tables II and III and Fig. 7). This would indicate that UGT1A7 may fulfill similar glucuronidation requirements as hepatic UGT1A9. UGT1A9, however, is absent in human gastric epithelium, where UGT1A7 may be the predominant UGT protein. Regulation of UGT1A7 and UGT1A10 in gastric epithelium, and UGT1A8, UGT1A9, and UGT1A10 in colonic epithelium possibly demonstrates gene control designed to lead to a specialized, tissue-specific profile of physiological glucuronidation. Further analysis will be required to characterize the full range of substrates for these proteins, in particular UGT1A8, and to elucidate their specific contribution to tissue-specific, extrahepatic glucuronidation.

In summary, this study documents the differential expression of the human *UGT1A* locus in hepatic and extrahepatic tissues of the gastrointestinal tract. A unique and diverse pattern of UGT1A regulation is identified in human colon leading to the discovery of a novel extrahepatic UGT1A8 protein. The tissue-specific regulation of UGT1A7 in gastric and UGT1A8 in colonic epithelium together with the identification of all transcripts of the related *UGT1A7-1A10* gene cluster in the gastrointestinal tract conclusively demonstrates a highly specific differential regulation of extrahepatic UGT1A expression.

REFERENCES

- Dutton, G. J. (1980) *Glucuronidation of Drugs and Other Compounds*, CRC Press, Boca Raton, FL
- Ritter, J. K., Chen, F., Sheen, Y. Y., Tran, H. M., Kimura, S., Yeatman, M. T., and Owens, I. S. (1992) *J. Biol. Chem.* **267**, 3257-3261
- Ebner, T., Rimmel, R. P., and Burchell, B. (1993) *Mol. Pharmacol.* **43**, 649-654
- Green, M. D. and Tephly, T. R. (1996) *Drug Metab. Dispos.* **24**, 356-363
- Kim, P. M., Winn, L. M., Parman, T., and Wells, P. G. (1997) *J. Pharmacol. Exp. Ther.* **280**, 200-209
- Mojarrabi, B., Butler, R., and Mackenzie, P. I. (1996) *Biochem. Biophys. Res. Commun.* **225**, 785-790
- Ritter, J. K., Crawford, J. M., and Owens, I. S. (1991) *J. Biol. Chem.* **266**, 1043-1047
- Owens, I. S. and Ritter, J. K. (1995) *Prog. Nucleic Acid Res.* **51**, 306-338
- Monaghan, G., Clarke, D. J., Povey, S., See, C. G., Boxer, M., and Burchell, B. (1994) *Genomics* **23**, 496-499
- Chowdhury, J. R., Chowdhury, N. R., Wolkoff, A. W., and Arias, J. M. (1994) in *The Liver: Biology and Pathobiology* (Arias, I. M., Boyer, J. L., Fausto, W. B., Jakoby, D. A., Schachter, D. A., and Schafritz, D. A. eds) pp. 471-504, Raven Press, New York
- Ritter, J. K., Yeatman, M. T., Ferreira, P., and Owens, I. S. (1992) *J. Clin. Invest.* **90**, 150-155
- Moghribi, N., Clarke, D. J., Boxer, M., and Burchell, B. (1993) *Genomics* **18**, 171-173
- Seppen, J., Bosma, P. J., Goldhoorn, B. G., Bakker, C. T., Chowdhury, J. R., Chowdhury, N. R., Jansen, P. L., and Oude Elferink, R. P. (1994) *J. Clin. Invest.* **94**, 2385-2391
- Philipp, T., Durazzo, M., Trautwein, C., Alex, B., Straub, P., Lamb, J. G., Johnson, E. F., Tukey, R. H., and Manns, M. P. (1994) *Lancet* **344**, 578-581
- Strassburg, C. P., Obermayer-Straub, P., Alex, B., Durazzo, M., Rizzetto, M., Tukey, R. H., and Manns, M. P. (1996) *Gastroenterology* **111**, 1576-1586
- Strassburg, C. P., Manns, M. P., and Tukey, R. H. (1997) *Cancer Res.* **57**, 2979-2985
- Wooster, R., Sutherland, L., Ebner, T., Clarke, D., Da Cruz e Silva, O., and Burchell, B. (1991) *Biochem. J.* **278**, 465-469
- Harding, D., Fournel-Gigleux, S., Jackson, M. R., and Burchell, B. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8381-8385
- Strassburg, C. P., Oldhafer, K., Manns, M. P., and Tukey, R. H. (1997) *Mol. Pharmacol.* **52**, 212-220
- Wei, Q., Xu, X., Cheng, L., Legerski, R. J., and Ali-Osman, F. (1995) *Cancer Res.* **55**, 5025-5029
- Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254
- Nguyen, N. and Tukey, R. H. (1997) *Drug Metab. Dispos.* **25**, 745-749
- Li, Q., Lou, X., Peyronneau, M., Straub, P. O., and Tukey, R. H. (1997) *J. Biol. Chem.* **272**, 3272-3279
- Von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99-105
- Ebner, T. and Burchell, B. (1993) *Drug Metab. Dispos.* **21**, 50-55
- Burchell, B., Nebert, D. W., Nelson, D. R., Bock, K. W., Iyanagi, T., Jansen, P. M., Lancet, D., Mulder, G. J., Chowdhury, J. R., Siest, G., Tephly, T. R., and Mackenzie, P. I. (1991) *DNA Cell Biol.* **10**, 487-494